EAI Endorsed Transactions
on Pervasive Health and Technology
Research Article

Use of personal mobile technologies for peer-based assessment of stress: a systematic literature review

Antoine Bellanger[1], Igor Matias[1 2,*], Katarzyna Wac[1]

[1]Quality of Life Technologies Lab, University of Geneva, Carouge, Switzerland
[2]Cognitive Aging Lab, University of Geneva, Carouge, Switzerland

Abstract

The use of personal mobile technologies has grown in recent years, providing a method for collecting high-frequency and high-quality data on human behaviors and states, amongst the others, on stress levels. Mobile technologies can play a significant role in peer-based stress assessment, particularly in e-mental health and well-being. It is accessible, convenient, and reliable compared to traditional self-report methods, making it a popular choice for collecting data. This systematic literature review aimed to explore the use of mobile technologies for peer-based assessment of stress. We analyzed existing literature to understand how mobile technologies have been used to assess stress levels through peer feedback—from relatives, friends, or others with close and daily contact with the individual. The results of the review showed that mobile technologies have the potential to be a valuable tool for peer-based stress assessment, as they can provide real-time and convenient data collection. However, although its popularity has grown in recent years, it is worth noting that the use of paper and pen questionnaires has remained prevalent in peer-based stress assessment over the last decade. This indicates that there is still a need for further exploration and evaluation of the benefits and limitations of both methods.

Received on 30 September 2024; accepted on 16 December 2024; published on 19 March 2025

Keywords: Mobile technologies, peers, stress, systematic literature review, behaviors, human states

doi:10.4108/eetpht.11.8941

1. Introduction

Stress is defined as "an unhealthy imbalance in a person's psychological and physiological state" [1]. It is a complex response to environmental demands that require an individual to adapt or cope with changes in their surroundings and internal state. It is a common state that can trigger various situations, including work-related challenges, social interactions, financial difficulties, and health concerns [2].

Stress can manifest in different ways, including cognitive, emotional, and physiological responses. Cognitive responses refer to how individuals perceive and interpret stressors, which can lead to negative thoughts and beliefs about their ability to cope effectively [3]. Emotional responses involve feelings of anxiety, fear, or sadness that may arise from stressors [4]. Physiological responses refer to changes in the body, such as increased heart rate [5], blood pressure, and cortisol levels, that prepare the body to cope with the stressor [6], [7].

Individuals also differ in their ability to cope with stressors, which personality traits and behaviors can influence. However, the relationship between stress, behaviors, and traits is complex and bidirectional. Stress can also lead to changes in behaviors and personality traits. For example, chronic stress can lead to anxiety or depression, altering an individual's personality traits (long term) and behaviors (short and long term) [8].

Personality traits and behaviors are two distinct but related concepts in psychology. Traits refer to enduring patterns of thoughts, feelings, and behaviors that characterize an individual's way of functioning in various situations. In contrast, behaviors refer to the

*Corresponding author. Email: igor.matias@unige.ch

specific actions, reactions, or responses an individual exhibits in each case [9]. Personality traits are relatively stable and consistent over time and across situations. They are believed to be influenced by a combination of genetic, environmental, and cultural factors. Behaviors, on the other hand, are more context-specific and variable. They are influenced by situational factors, such as social norms, environmental cues, and personal goals [10]. Figure 1 depicts the influences between stress, traits, and behaviors.

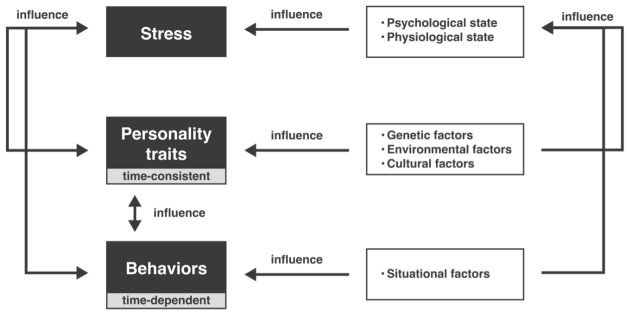

Figure 1. Influence existing between stress, personality traits, and behaviors.

Traits like being more neurotic, anxious, or perfectionistic lead an individual to be more susceptible to stress and have greater difficulty coping with it than those who are more resilient or have a positive outlook on life. Additionally, coping behaviors such as exercise, social support, and relaxation techniques can help individuals manage stress more effectively. In contrast, unhealthy behaviors such as substance abuse, overeating, and avoidance can exacerbate the negative effects of stress [8] [11].

Stress affects individuals of all ages. Some age groups, like adolescents, are especially vulnerable to it, regardless of their activities [12]. For example, stress can particularly affect students whose emotions are closely linked to their ability to learn and perform well academically [13]. According to Adesola et al. [14], higher anxiety during test evaluations leads to generally worst academic success. Gender differences were also found by Sverdlik et al. [15], with women being more impacted than male students.

Within the surrounding sources of stress, pandemics like COVID-19 are identified as significant ones. The stress caused by those can be compounded by the numerous challenges individuals face during the outbreak. Leung et al. [16] reported that pandemics generate an increase in mental disorders. Internal states and behaviors, for example, related to smartphone addiction, have also been linked to increased stress and negative mental health outcomes such as depression and anxiety [17].

Subjective human states such as stress are traditionally evaluated using self-report methods denoted as "patient-reported outcomes" (PROs, as defined by Mayo et al. [18]). Even though the term patient does not always apply to the assessed participants in the studies included in this review, we refer to PROs using that same designation to match the definition of PRO previously introduced.

PROs include self-reports before making any interpretation. PROs are affected by reporting bias when the individuals hide the real answers for privacy, social, or other reasons. That bias can be found in our everyday life. For example, students can grade themselves harsher than their tutors or peers, as presented by Stefani et al. [19]. The opposite happens when assessing stress levels, with patients underestimating their stress severity compared to their peers. In that case, we leverage peer reporting, known to reduce bias and be more reliable than PROs, as demonstrated by Hogset et al. [20]. Peer-reported outcomes (PeerROs) capture data on individuals' states involving their peers. Those can include, for example, colleagues and family members [8] [9]. In clinical settings, it can imply trained healthcare practitioner (eg. a nurse).

A popular way to collect momentary PROs and PeerROs is through ecological momentary assessment (EMA), a method of reporting data on states, thoughts, and behaviors as they occur in real-time and natural environments. It involves regularly collecting data in various ways, such as short or long questionnaires. Even though EMA was since recently typically implied burdensome self-reports on a paper, recent personal technological developments have been made to decrease this burden [21]. That allows the individual to answer questionnaires easier and at any moment. As shown by Runyan et al. [22], using new technologies also improves the quality of the collected data and the easiness of processing larger data sets [11].

Mobile devices such as conventional smartphones are omnipresent in our daily life. Dey et al. [23] found those near us almost 90% of the day. They contain sensors that can provide new information enriching EMA, such as sleep duration, physical activity, or phone usage [24]. Manea et al. [25] reported strong correlations between PROs and technology-reported outcomes (TechROs) collected with a Fitbit Charge

2®[1] device. Those correlations were found in different domains, such as physical activity, sleep, anxiety, and depression. But smartphones can also act as a mere data collection method for EMA by increasing the frequency and simplicity with which PROs and PeerROs can be obtained. Specifically addressing stress monitoring, a higher granularity of EMA means a more accurate understanding of the individual state, as states and emotions can vary too much during the day.

This research aims to understand current practices and implications of using personal technologies to acquire PeerRO. Even though a scoping literature review could achieve that goal, this work relies instead on the systematic literature review approach. It does so because, despite the literature search criteria being broad enough for a scoping, our conclusions and discussions focus solely on analyzing the use of personal mobile technologies, thus funneling the information into a specific literature area. This systematic literature review provides an in-depth analysis and discussion of the usage of portable technology for peer-based stress reporting regarding frequency, quality, advantages, and disadvantages. Thus, we first screen for all recent peer-based stress assessment research results. Secondarily, we specifically look at technology usage (or its non-usage) while conducting such an assessment. This review has the following goals:

1. Provide an overview and discussion of all the peer-based stress assessments conducted over the last decade (PROs and PeerROs).

2. Identify the methods, including validated (or not) questionnaires, employed in such an assessment.

3. Pinpoint examples of research utilizing technology during the stress assessment to enable similar and more advanced research.

4. Discuss the advantages and disadvantages of using technology while assessing stress and other states with PROs and PeerROs.

The remainder of this article is organized as follows. Section 2 describes the method designed for eligibility selection, the extraction of information from the included articles, and the research questions. Section 3 presents the results of the search using summary tables. In Section 4, we discuss the results and answer the research questions. Lastly, Section 5 contains the conclusions and recommendations from this systematic literature review.

[1] Fitbit Charge 2 is a registered trademark of Fitbit, Inc.

2. Methods

This systematic literature review was conducted following the recommendations by Silva et al. [26] and according to the *Preferred Reporting Items for Systematic Reviews and Meta-Analysis* (PRISMA) statement [27].

2.1. Search Strategy

A literature search was done on four databases: IEEE Xplore, ACM Digital Library, PubMed, and ScienceDirect. The search included results between the 1st of January 2012 and the 22nd of August 2022. The first two databases were chosen due to their inclusion of high-quality articles in the computer science field. PubMed was selected for its inclusion of articles in the areas of biomedicine and medicine. ScienceDirect was included for containing thousands of books and journals in many fields. The search was done according to the titles and abstracts and the keywords presented below. The list of references of the selected articles was also screened for the inclusion of any additional relevant articles. The keywords used on all four databases were: (peer) AND (assessment OR measure) AND (stress)

2.2. Study selection

We first reviewed the titles and the abstracts of the eligible articles according to our search criteria using *Rayyan QCRI* [28]. The key works aimed to include research articles on stress assessment using peers. That is, for example, colleagues, family members, and medical staff with close contact with the individual who can evaluate their stress levels. The search focused solely on stress disregarding peer-based assessment of other pathologies like depression or anxiety.

During the first stage, a single reviewer analyzed the title and the abstract of the papers and included, excluded, or doubted the article. The included and doubtful articles were then analyzed by two reviewers in the second stage. From the introduction and the conclusion, a decision was made to include or exclude the paper. In the case where the two reviewers disagreed on a choice, a diagonal reading of the paper was made by them to agree on a common decision.

The inclusion and exclusion criteria are summarized in Table 1. All articles not excluded were then analyzed based on their introductions and conclusions. This was made to evaluate their relevancy (or not) to our search. Finally, the eligible articles were evaluated based on their full text.

2.3. Extraction of study characteristics

We extracted characteristics from the selected studies to analyze the experimental setup, the methods employed

Table 1. Inclusion and exclusion criteria used in this review [29]

Type	Inclusion	Exclusion
Date	Between 1 January 2012 and 22 August 2022	Before 1 January 2012
Geographic location of study	All	None
Language	English	Any other languages
Study design	All	None
Participants	All	None
PeerROs used	At least stress levels	None
Peer review	Journal and conference	All the others
Type of publication	Journal and conference	All the others

in the data collection, the questionnaires used, and the analysis made by the authors:

- Article Information: defines the country where the study was conducted and the year, type, and place of publication.
- Participants: defines the participants (study subjects and their peers), the sample size and participants' age categories, and the method employed to recruit them.
- Questionnaires: defines which questionnaires were used in the experiments, namely which one was used for the stress peer-assessment, and the frequency of data collection. This only includes the questionnaires used as PeerRO, that is, only those answered by the peers. If the peer is a trained clinician, the PeerRO becomes ClinRO. [18]
- Data collection methods: defines the collection methods used to record the assessments.
- Analysis methods: defines how the authors analyzed the data.

2.4. Research questions

This review had the following research questions referring to the peer-based stress assessment:

Who are the peers, and which methods are used to recruit the peers?

Which questionnaires are employed, and at which frequency?

Which data collection methods are used to collect the peer-based stress data?

Which methods are used to analyze the collected data?

How is the stress assessment data used?

RQ1's goal was to list the types of peers and methods used to recruit the peers involved in the included studies. The answer to RQ1 is expected to help future research in the field during that initial step. RQ2 and RQ4 were defined to list the most employed questionnaires and review the most used analysis methods in the field. The motivation for RQ3 was to identify the data collection methods used for the peer-assessment of stress (e.g., using new technologies or regular paper-based questionnaires). Finally, RQ5 was motivated by a need to summarize how PeerRO's data can be used.

3. Results

The initial search results returned 784 unique articles after removing 71 duplicates. Following the review of the title and the abstract, according to the criteria defined in Table 1, we excluded 750 articles leaving us with a remaining 34 articles. When retrieving those articles, the full body of two of them could not be obtained (the authors were contacted, but there was no reply within 15 days). The following step was to read the introduction and the conclusion of the 32 retrieved articles. From those, 16 were excluded because they had no assessment of stress, nine because they did not have a stress assessment based on a peer's input (i.e., PeerRO), and one because it was not written in the English language. Ultimately, a total of six articles were included in this review. After screening their references for additional relevant work, we added one article to this systematic review. The flow diagram of the selection process of the articles is shown in Figure 2.

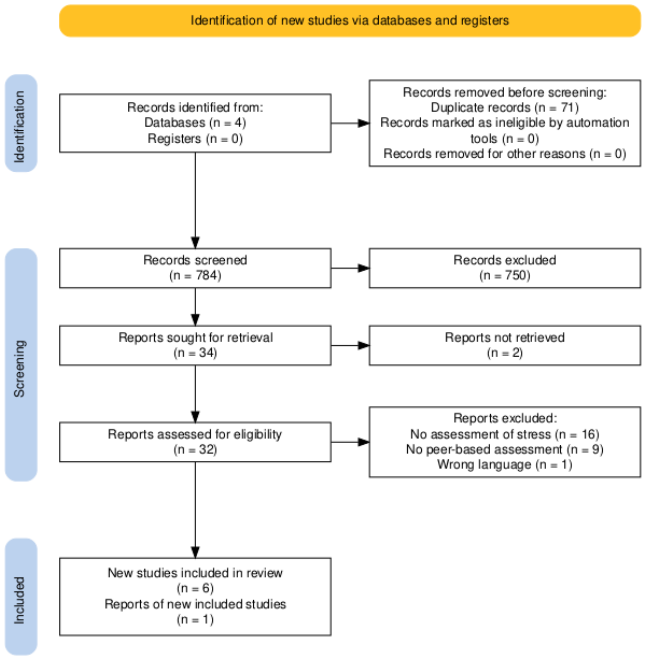

Figure 2. Flow diagram of identification and inclusion papers [30].

Table 2. Participants' information collected from the studies. (NR = Not Reported/Not Applicable)

Year	Study	Number of study participants	Age of the patients	Type and number of peers	Age peers
2010 (study until 2006)	Kromm et al. [33]	384	45.9 (9.0) years old	275 spouses, 127 friends	44.2 (9.1), 43.7 (11.7)
2012 (study until 2007)	Gunnarsdottir et al. [34]	91	11.02 (1.42) years old	91 parents	40.34 (5.31)
2015	Lydon et al. [35]	57	10.89 (4.25) years old	NR	NR
2018	Berrocal et al. [1]	NR	NR	NR	NR
2019	Palmer et al. [31]	22	4 to 8 years old	22 parents, 22 teachers	NR
2020 (study in 2018)	Berrocal et al. [36]	13	Above 18 years old	20	Above 18 years old
2020	Berrocal et al. [37]	23	Above 18 years old	27	Above 18 years old

3.1. Eligibility of studies

Although all the included studies respected our inclusion and exclusion criteria, it is useful to clarify the selection process for one of them. While the study done by Palmer et al. [31] did not focus primarily on the peer-based assessment of stress, parents and teachers had to fill out the Aberrant Behavior Checklist [32] which indirectly included stress. Therefore, a peer-based assessment of stress was made but indirectly. For that reason, that article was included in this review.

3.2. Source of evidence

Out of the seven included articles, four (57%) were published in a medicine journal, two (29%) in a computer science conference proceeding, and one (14%) in a health informatics journal.

3.3. Study participants and design

The included studies used a variety of study participants ranging from children to adults, while the peers were always above 18 years old. Both the number and age of the study participants are not reported by one study (14.28%). In two studies (28.57%), the number of peers is missing, and in three studies (42.85%), the age of the peers is not reported. The participants' information is presented in Table 2.

3.4. Data collected from selected studies

Although it was necessary to gather as much information as possible on the selected studies while performing the quality synthesis process, some articles contained data not comparable with any other. For that reason, this section includes only the data similar between at least two articles included in this review.

Methods used to recruit the peers. Table 4 presents the methods used to recruit peers from all the selected studies. By doing so, it exemplifies and compares the multiple ways to look for peers when performing research like the one focused on by this systematic literature review.

Table 3. Methods used to recruit the peers. (NR = Not Reported/Not Applicable)

Study	Method(s) to recruit the peers
Kromm et al. [33]	Leadership seminars, Meetings with German and Austrian companies, Contacts with German staff working at NATO
Gunnarsdottir et al. [34]	School nurses contacting parents of obese children
Lydon et al. [35]	NR
Berrocal et al. [1]	NR
Palmer et al. [31]	Referral via local autism diagnostic teams, Education professionals, Support groups, Consented databases, Self-referral
Berrocal et al. [36]	Flyers, Emails, Word of mouth around university
Berrocal et al. [37]	Flyers, Emails, Word of mouth and advertisement around university

Table 4. Questionnaires employed, frequency of data collection, and data collection method used.

Study	Questionnaires employed	Frequency of collection	Data collection method
Kromm et al. [33]	*Trier Inventory for Assessment of Chronic Stress* ("TICS" - PRO) [38]	Once	Pen and paper
Gunnarsdottir et al. [34]	*Children's Depression Inventory* ("CDI" - PRO) [39], *Multidimensional Scale for Children* ("MASC" - PRO) [40], *Strengths and Difficulties Questionnaires* ("SDQ" - PeerRO) [41]	Once	Pen and paper
Lydon et al. [35]	*Stress survey schedule for individuals with autism and other developmental disabilities* ("SSS" - ClinRO) [42]	Once	Pen and paper
Berrocal et al. [1]	*Perceived Stress Scale* ("PSS" - PRO) [43], *Social Desirability Scale* ("SDS" - PRO) [44]	Eight times per day for 28 days	Mobile application
Palmer et al. [31]	*ABC irritability and hyperactivity* ("ABC" - PeerRO) [32]	Twice, at baseline and post-intervention (6 to 8 months apart)	Pen and paper
Berrocal et al. [36]	Subjective single-item questions (PeerRO)	Eight times per day for 28 days	Mobile application
Berrocal et al. [37]	*Perceived Stress Scale* ("PSS" - PRO) [43], *Social Desirability Scale* ("SDS" - PRO) [44]	Eight times per day for 28 days	Mobile application

When analyzing the selected articles, the authors did not specify their methods in two (28.57%) studies. Two others (28.57%) recruited participants through their university with flyers, emails, advertisements, and word of mouth. One (14.29%) study focused on executives ("German military officers serving at the North Atlantic Treaty Organization" and senior employees as in [33]) and recruited them in seminars and companies. Another one (14.29%) used the support of local autism diagnostic teams, education professionals, support groups, consented databases, and self-referral.

Questionnaires were employed and data collection methods. Table 4 shows, for each study, which questionnaires the peers had to fill out at each frequency and with which data collection method. The data presented in this table is solely focused on the peer-based assessment part of the study (PeerROs, even if done with PROs repurposed for peer-based assessment). Additionally, Table 5 does not include any information on the assessment of any other state, as some of the studies required the peers to fill out other questionnaires.

Table 5. Analysis methods and usage of the peer-based stress assessment data.

Study	Analysis	Usage of data
Kromm et al. [33]	Cronbach's alpha, Dependent t-tests, Exploratory factor analyses, Oblique rotation	Comparison between PeerRO and PRO
Gunnarsdottir et al. [34]	Standard multiple regression	Comparison between PeerRO and PRO
Lydon et al. [35]	Correlations, Hierarchical regressions	Comparison between PRO and ClinRO
Berrocal et al. [1]	Model to classify stress, Mixed-methods (qualitative and quantitative) analysis	Comparison between PeerRO and PRO
Palmer et al. [31]	*Analysis of covariance* (ANCOVA)	Direct assessment of stress from the peer reporting
Berrocal et al. [36]	Machine-learning classification algorithms	Comparison between PeerRO and PRO
Berrocal et al. [37]	Visualization of raw datasets, Spearman ranked correlation, Wilcoxon signed-ranked test	Comparison between PeerRO and PRO

Regarding the questionnaires used to obtain PeerROs of stress, every study used a different one except for two studies ([1] and [13], both by Berrocal et al.). Each study had at least one questionnaire about stress or a reaction caused by stress (e.g., irritability). The frequency of data collection was mainly only once (three of the seven included articles—42.86%). If we cluster all the included studies between low (data collected once or twice across the study) and high (more than three times), only three out of the seven performed high-frequency PeerROs collection, all from the same author (Berrocal et al.). When collecting data from the individuals and their peers, four out of the seven studies (57.14%) collected data using paper and pencil ([31], [33]–[35]). The three other studies collected data using a mobile application and were all from the same author ([1], [36], [37]). If analyzed together, the high frequency of data collection matches the usage of mobile applications as the data collection method.

Analysis and usage of stress data. The analysis methods used and the objective of using PeerROs in each study are described in Table 5. It serves as a comparison tool between the possible ways to use PeerROs of stress and the respective adequate analysis methods.

For the data analysis performed by the researchers, we also noticed that all seven studies employed different techniques to obtain results. Most studies (71.43%) followed statistical approaches, while only two (28.57%) used machine-learning methods. Regarding the usage of PeerROs of stress, five of the seven studies (71.43%) used the stress assessment data to compare the results between the peer-based and the study participant-based assessments of stress. Another (14.29%) used the data to compare the peer-based stress assessment with clinician-based stress (cortisol levels). Finally, one last study (14.29%) used it to do a direct assessment of stress from peer reporting.

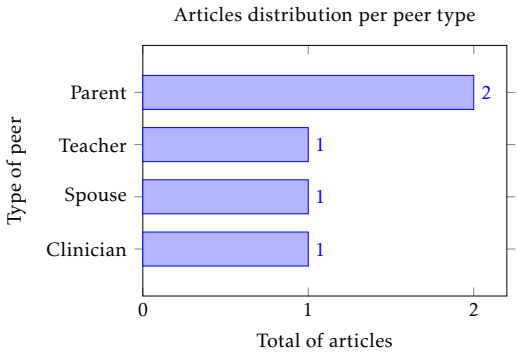

Figure 3. Diagram representing the number of times a method was used.

4. Discussion

Given the aim of this systematic literature review is the identification and exploration of the articles involving peer-based assessment of stress levels, this section discusses each of the research questions previously introduced.

4.1. Who are the peers, and which methods were used to recruit the peers? (RQ1)

In the seven articles, peers are not reported in two of them or not categorized in two others. In one article, peers are parents and teachers, and in another, they are only the parents. Finally in the last article, they are the spouse and friends of the study participants.

Two articles did not provide information on how they recruited the peers, while the other five mentioned one or more methods in different categories. We grouped the methods to recruit peers into three categories: word-of-mouth, online and physical advertisement. Two of the seven articles analyzed did not provide information on the recruitment methods used. However, the remaining five articles mentioned one or more methods from the three categories.

The word-of-mouth method was the most used in the articles, with five mentioning its use. This method involved local specialized teams, support groups, or word-of-mouth referrals from other participants. This approach highlights the importance of building a supportive community in peer-based assessments and the role of trust in the recruitment process.

Three articles used online methods to recruit peers. This involved using online services such as databases and emailing potential participants. This approach reflects the current trend toward online communication and the importance of using technology to reach a wider audience.

Physical advertisement methods, such as flyers and physical announcements, were used in two articles. This approach highlights the importance of reaching

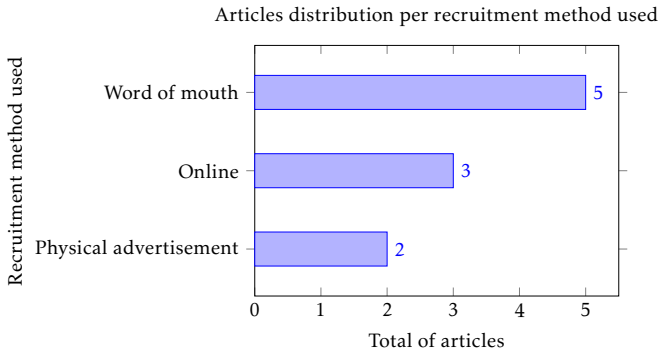

Figure 4. Diagram representing the number of times a method was used.

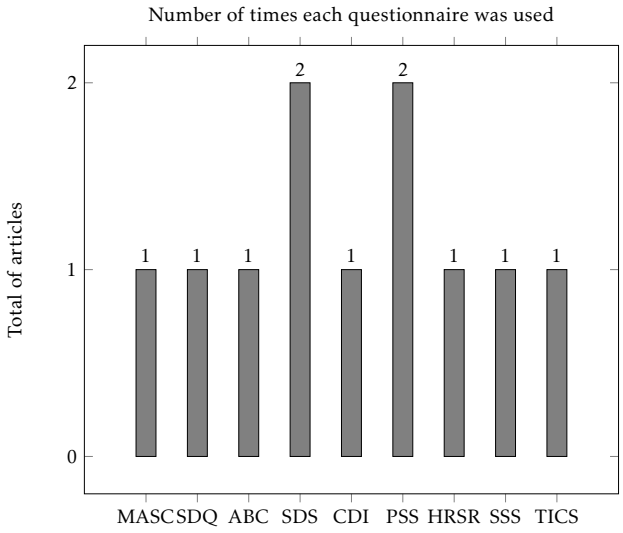

Figure 5. Diagram representing the number of articles in which a specific questionnaire was used.

participants who may not have access to technology or prefer more traditional communication methods.

In conclusion, the methods used to recruit peers using mobile technologies for peer- based assessment of stress varied across articles, with word-of-mouth being the most used method. That suggests that a combination of different recruitment methods may be necessary to reach a diverse group of participants. In the future, it would be beneficial to continue exploring new and innovative methods for recruiting peers, including social media, gamification, and other digital technologies. Additionally, these methods should be evaluated regarding their impact on the overall success of the peer- based assessment and the level of participation and engagement of peers.

4.2. Which questionnaires are employed, and at which frequency? (RQ2)

The authors of the selected articles utilized various questionnaires, which can be broadly categorized into four groups depending on the main focus of each tool, including (1) anxiety, (2) depression, and (3) stress states, and (4) behaviors and personality traits. The three first did include stress-related questions and are included in this RQ for that same reason.

We found that only one study used the *Multidimensional Anxiety Scale for Children* (MASC) [40] to evaluate anxiety levels among study participants. This may reveal a lack of measurement of anxiety levels in the peer-based assessment of stress. Further research is needed to establish a commonly used questionnaire to evaluate anxiety levels in such studies.

For depression assessment (but still including indirect stress assessment), a single questionnaire was utilized, namely the *Children's Depression Inventory* (CDI) [39]. Again, the frequency of usage was low, with only one study using the questionnaire.

Regarding stress, a variety of questionnaires were used for purely stress assessment, including the *Trier Inventory for the Assessment of Chronic Stress* (TICS) [38], the *Stress survey schedule for individuals with autism and other developmental disabilities* (SSS) [42], the *Holmes and Rahe Social Readjustment Rating Scale* (RHSRRS) [45], and the *Perceived Stress Scale* (PSS) [43]. PSS was used twice, while the others were used only once. This highlights the need for further research to establish commonly used questionnaires for assessing perceived stress levels using PeerROs. Despite purely stress assessment questionnaires being used by many researchers, the most popular type of questionnaire was behavior-based. That may be justifiable because most of the included articles focused on the stress levels of autistic children, mainly manifested by their conduct.

Finally, for the evaluation of behavior, a variety of questionnaires were used, including the *Strengths and Difficulties Questionnaire* (SDQ) [41], the *ABC irritability and hyperactivity* (ABC) [32] and the *Social Desirability Scale* (SDS) [44]. The frequency of usage of these questionnaires was low, with one study utilizing the first two (SDQ and ABC) and two studies using SDS.

Peers filled out nine questionnaires. Out of those, four (44.44%) were thought of for peer-reporting by design. The remaining five (55.56%) were made for the individual to answer directly (i.e., via a PRO) and adapted to be reported by a peer.

In conclusion, we highlight the need for standardized, validated questionnaires for the peer-based stress assessment. Further research is needed to establish commonly used questionnaires for evaluating anxiety, depression, and stress states, and behaviors, and personality traits, in such peer-based studies.

4.3. Which data collection methods are used to collect the peer-based stress data? (RQ3)

Some studies utilized physical questionnaires (pen and paper), and others opted for mobile applications. The use of mobile technologies as a method for collecting peer-based stress data has grown in recent years. This is indicated by the fact that three of the seven studies conducted in the literature review used mobile applications, all of which were recent (2019 and after). The author Berrocal et al. conducted all these studies, which might suggest that the use of mobile technologies in this field is still in its initial stages and is being explored by a few authors.

However, it is crucial to note that this does not necessarily mean that the use of mobile technologies for peer-based stress assessment is limited to one author, as other studies might have used those same and other methods and may not have been included in this systematic literature review.

In conclusion, the use of mobile technologies as a method for collecting peer-based stress data is on the rise and holds significant potential for future research in this field. The benefits of using mobile technologies include improved accessibility and the potential to increase the participation rate of individuals who might not otherwise have participated in traditional pen-and-paper questionnaires [46]. The advancement of technology has opened new avenues for researchers, and it is vital to continue exploring mobile technologies in peer-based stress assessment.

4.4. Which methods are used to analyze the collected data? (RQ4)

The analysis of collected data plays a crucial role in the validity and reliability of the results. The method used to analyze the data can greatly impact the results and conclusions drawn from the research. Two main methods were used to analyze the collected data in the seven articles, including statistical analysis and machine learning. In the studies that utilized statistical analysis, the authors applied correlations, regressions, and analysis of covariance to understand the relationship between different variables. These methods are well-established and commonly used in the field of statistics [47], allowing for examining relationships between variables and identifying any potential trends or patterns. Using these methods provided a systematic approach for the authors to examine the data and draw conclusions about the effectiveness of peer-based assessment of stress using mobile technologies.

However, it is also worth noting that statistical analysis has limitations, such as the assumption of linear relationships between variables [48]. Additionally, statistical analysis can only examine relationships between variables and may not be able to identify more complex patterns in the data.

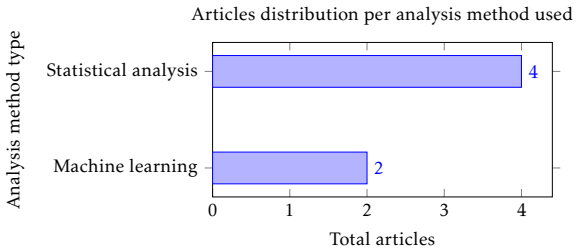

Figure 6. Diagram of the data analysis methods used by the authors in the seven papers.

In contrast, the studies that utilized machine learning employed a variety of algorithms, including Decision Tree, Random Forest, Extreme Gradient Boosting, and ZeroR. Machine learning methods can identify more complex data patterns and make predictions based on the identified patterns. These methods are particularly useful in fields where data sets can be large and complex, and relationships between variables may not be easily identifiable through traditional statistical methods [49].

In conclusion, the statistical analysis allowed for examining relationships between variables, while machine learning methods allowed for identifying more complex patterns in the data. However, it is important to keep in mind each method's limitations and consider using a combination of methods for a more robust data analysis. Moving forward, it may be worth exploring other advanced methods, such as deep learning or reinforcement learning to enhance further the analysis of collected data especially frequent, longitudinal data, where peer-based assessment of stress using mobile technologies.

4.5. How is the stress assessment data used? (RQ5)

Most studies used the stress PeerROs collection for direct comparison with the PROs of stress provided by the individuals themselves. That calibration with PROs is useful as it compares self-reported stress levels and those reported by peers. It offers valuable insights into the consistency and accuracy of self-reported stress levels, which is essential for diagnosing individuals or developing effective treatment plans and assessing their efficacy.

Additionally, a minority of studies used the stress PeerROs collection for calibration with other outcomes, such as the clinician-based assessment of stress levels. Another usage of those PeerROs was the assessment of a study participants' stress levels without their input, that is, by using only the peer-based assessment. This highlights the versatility of peer-based stress assessment and how it can be used differently to

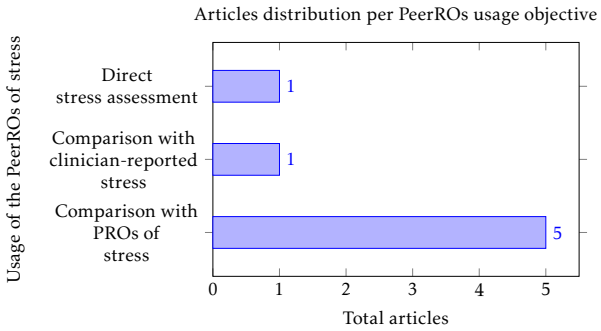

Figure 7. Diagram of the usage types of the PeerROs for stress.

provide valuable insights into the stress levels of individuals.

In conclusion, the versatility of peer-based stress assessment highlights its potential for future research in this field and its usefulness in improving the understanding of stress levels.

5. Conclusions

This systematic literature review presents and summarizes the current work on the peer-based assessment of perceived stress in individuals, particularly analyzing the usage of mobile technologies as a data collection method.

The more common source of peers recruited was the study participants' household, which includes parents (two out of seven articles) and partners (one article). However, it is important to note that the only two studies using parents as peers studied the levels of stress among children. That can possibly explain why this type of peer (parents) was not reportedly used in any other study included in this review.

When focusing on the strategies for recruiting peers for stress assessment studies, we found that the methods used varied across studies, with word-of-mouth being the most used method. Using a combination of different recruitment methods may be necessary to reach a diverse group of participants. It would be beneficial to continue exploring new and innovative ways to recruit peers, including social media, gamification, and other digital technologies.

We also found a need for standardized questionnaires for the peer-based stress assessment. Further research is needed to establish commonly used peer-based questionnaires for evaluating anxiety, depression, and stress states, and behaviors, and personality traits in such studies to ensure the findings' validity and reliability at the frequency corresponding to changing the phenomenon being evaluated. Additionally, we found a prevalent low-frequency of stress assessment, which can limit the results of such research. As stress and other human states are highly dynamic and depend on multiple time-dependent factors, its low evaluation frequency means a high probability of missing important information periods.

The use of mobile technologies as a method for collecting peer-based stress data has grown in recent years, as indicated by the fact that three of the seven studies conducted in the literature review used mobile applications. This shift is likely due to its accessibility and convenience compared to traditional methods like paper and pen questionnaires. The benefits of using mobile technology include improved accessibility and the potential to increase the participation rate of individuals who might not otherwise have access to peer-based assessments. However, it is also worth noting that despite the growing popularity of mobile technology, the use of paper and pen questionnaires remains prevalent in peer-based stress assessment. This indicates that there is still a need for further exploration and evaluation of the benefits and limitations of both methods.

As a suggestion for future research in this domain, we strongly recommend the authors report some details of the studies further aside from the ones described in Section 2.3. Namely, we encourage the addition of (1) the duration of the recruitment period and (2) the percentage of accepted and not accepted participants recruited, which will enable a better understanding of the success of the recruitment tools used, and (3) the limitations and discussion of possible alternatives of the methods used during the data analysis phase of the studies. Further limitation of this study is that it only takes into account papers until 2022, as it took time to code the paper and prepare this review.

Overall, the findings of this systematic literature review indicate that mobile technologies hold significant potential for peer-based assessment of stress, and further research in this field is needed. The use of mobile technologies should be evaluated in terms of its impact on the overall validity of the peer-based assessment and the level of participation and engagement of peers. The future of peer-based stress assessment using mobile technologies holds great promise and is an area that deserves continued attention and exploration.

Acknowledgements

This study was supported by the AAL Guardian (2019-6-120-CP), swissuniversities P-13 AGE-INT, UCPH Data+ AI@CARE Project and Centre Universitaire d'Informatique of the University of Geneva.

References

[1] Berrocal, A. and Wac, K. (2018) Peer-ceived well-being: Exploring the value of peers for human stress assessment in-situ. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International*

Symposium on Pervasive and Ubiquitous Computing and Wearable Computers, UbiComp '18 (New York, NY, USA: Association for Computing Machinery): 492–497. doi:10.1145/3267305.3267319.

[2] Folkman, S. (2013) *Stress: Appraisal and Coping* (New York, NY: Springer New York), 1913–1915. doi:10.1007/978-1-4419-1005-9_215.

[3] Emotion and adaptation., https://psycnet.apa.org/record/1991-98760-000. [Accessed 02-03-2023].

[4] Parrott, W.G. (2001) Implications of dysfunctional emotions for understanding how emotions function. *Review of General Psychology* **5**(3): 180–186. doi:10.1037/1089-2680.5.3.180.

[5] Matias, I., Daza, E.J. and Wac, K. (2022) What possibly affects nighttime heart rate? conclusions from n-of-1 observational data. *DIGITAL HEALTH* **8**: 20552076221120725. doi:10.1177/20552076221120725. PMID: 36046637.

[6] Segerstrom, S.C. and Miller, G.E. (2004) Psychological stress and the human immune system: A meta-analytic study of 30 years of inquiry. *Psychological Bulletin* **130**: 601–630. doi:10.1037/0033-2909.130.4.601.

[7] Brosschot, J.F., Gerin, W. and Thayer, J.F. (2006) The perseverative cognition hypothesis: A review of worry, prolonged stress-related physiological activation, and health. *Journal of Psychosomatic Research* **60**(2): 113–124. doi:10.1016/j.jpsychores.2005.06.074.

[8] Connor-Smith, J.K. and Flachsbart, C. (2007) Relations between personality and coping: A meta-analysis. *Journal of Personality and Social Psychology* **93**: 1080–1107. doi:10.1037/0022-3514.93.6.1080.

[9] Roberts, B.W. and Mroczek, D. (2008) Personality trait change in adulthood. *Current Directions in Psychological Science* **17**(1): 31–35. doi:10.1111/j.1467-8721.2008.00543.x. PMID: 19756219.

[10] Funder, D.C. and Colvin, C.R. (1991) Explorations in behavioral consistency: Properties of persons, situations, and behaviors. *Journal of Personality and Social Psychology* **60**: 773–794. doi:10.1037/0022-3514.60.5.773.

[11] Gross, J.J. (2013) Emotion regulation: Taking stock and moving forward. *Emotion* **13**: 359–365. doi:10.1037/A0032135.

[12] Arun, P., Garg, R. and Chavan, B. (2017) Stress and suicidal ideation among adolescents having academic difficulty. *Industrial Psychiatry Journal* **26**: 64. doi:10.4103/IPJ.IPJ_5_17.

[13] Reinhard Pekrun, Thomas Goetz, W.T. and Perry, R.P. (2002) Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist* **37**(2): 91–105. doi:10.1207/S15326985EP3702_4.

[14] Adesola, S.A., Li, Y. and Liu, X. (2019) Effect of emotions on students learning strategies. In *Proceedings of the 2019 8th International Conference on Educational and Information Technology*, ICEIT 2019 (New York, NY, USA: Association for Computing Machinery): 153–156. doi:10.1145/3318396.3318408.

[15] Anna Sverdlik, N.C.H. and Vallerand, R.J. (2023) Doctoral students and covid-19: exploring challenges, academic progress, and well-being. *Educational Psychology* **43**(5): 545–560. doi:10.1080/01443410.2022.2091749.

[16] Leung, C.M., Ho, M.K., Bharwani, A.A., Cogo-Moreira, H., Wang, Y., Chow, M.S., Fan, X. *et al.* (2022) Mental disorders following covid-19 and other epidemics: a systematic review and meta-analysis. *Translational Psychiatry 2022 12:1* **12**: 1–12. doi:10.1038/s41398-022-01946-6.

[17] Zhao, L. and Hounnaklang, N. (2022) Association between smartphone addiction and mental health during the covid-19 pandemic 2021 among inner mongolia medical university students, china. In *Proceedings of the 2021 International Conference on Intelligent Medicine and Health*, ICIMH '21 (New York, NY, USA: Association for Computing Machinery): 115–124. doi:10.1145/3484377.3487040.

[18] Mayo, N.E., Figueiredo, S., Ahmed, S. and Bartlett, S.J. (2017) Montreal accord on patient-reported outcomes (pros) use series – paper 2: terminology proposed to measure what matters in health. *Journal of Clinical Epidemiology* **89**: 119–124. doi:10.1016/j.jclinepi.2017.04.013.

[19] Stefani, L.A. (1994) Peer, self and tutor assessment: Relative reliabilities. *Studies in Higher Education* **19**(1): 69–75. doi:10.1080/03075079412331382153.

[20] Hogset, H. and Barrett, C.B. (2010) Social learning, social influence, and projection bias: A caution on inferences based on proxy reporting of peer behavior **58**: 563–589. doi:10.1086/650424.

[21] Doherty, K., Balaskas, A. and Doherty, G. (2020) The Design of Ecological Momentary Assessment Technologies. *Interacting with Computers* **32**(3): 257–278. doi:10.1093/iwcomp/iwaa019.

[22] Runyan, J.D. and Steinke, E.G. (2015) Virtues, ecological momentary assessment/intervention and smartphone technology. *Frontiers in Psychology* **6**. doi:10.3389/fpsyg.2015.00481.

[23] Dey, A.K., Wac, K., Ferreira, D., Tassini, K., Hong, J.H. and Ramos, J. (2011) Getting closer: an empirical investigation of the proximity of user to their smart phones. In *Proceedings of the 13th International Conference on Ubiquitous Computing*, UbiComp '11 (New York, NY, USA: Association for Computing Machinery): 163–172. doi:10.1145/2030112.2030135.

[24] Wang, W., Nepal, S., Huckins, J.F., Hernandez, L., Vojdanovski, V., Mack, D., Plomp, J. *et al.* (2022) First-gen lens: Assessing mental health of first-generation students across their first year at college using mobile sensing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **6**(2). doi:10.1145/3543194.

[25] Manea, V. and Wac, K. (2020) Co-calibrating physical and psychological outcomes and consumer wearable activity outcomes in older adults: An evaluation of the coqol method. *Journal of Personalized Medicine* **10**(4). doi:10.3390/jpm10040203.

[26] Weidt, F. and Silva, R. (2006) Book review **11**: 244–245. doi:10.1027/1016-9040.11.3.244.

[27] (2009) Preferred reporting items for systematic reviews and meta-analyses: The prisma statement. *Annals of Internal Medicine* **151**(4): 264–269. doi:10.7326/0003-4819-151-4-200908180-00135. PMID: 19622511.

[28] Ouzzani, M., Hammady, H., Fedorowicz, Z. and Elmagarmid, A. (2016) Rayyan-a web and mobile app for systematic reviews. *Systematic Reviews* **5**: 1–10.

doi:10.1186/S13643-016-0384-4.

[29] Condron, P., Library guides: Systematic reviews: 5: Inclusion and exclusion criteria, https://unimelb.libguides.com/sysrev/inclusion-exclusion-criteria. Accessed: Nov. 22, 2022.

[30] Prisma flow diagram, https://prisma-statement.org/prismastatement/flowdiagram.aspx. Accessed: Nov. 22, 2022.

[31] Palmer, M., Tarver, J., Perez, J.P., Cawthorne, T., Romeo, R., Stringer, D., Hallett, V. *et al.* (2019) A novel group parenting intervention to reduce emotional and behavioural difficulties in young autistic children: protocol for the autism spectrum treatment and resilience pilot randomised controlled trial. *BMJ Open* **9**: e029959. doi:10.1136/BMJOPEN-2019-029959.

[32] Aman, M.G., Singh, N.N., Stewart, A.W. and Field, C. (1985) The aberrant behavior checklist: a behavior rating scale for the assessment of treatment effects. *American journal of mental deficiency* **89**(5): 485–491.

[33] Kromm, W., Gadinger, M.C. and Schneider, S. (2010) Peer ratings of chronic stress: can spouses and friends provide reliable and valid assessments of a target person's level of chronic stress? *Stress and Health* **26**: 292–303. doi:10.1002/SMI.1297.

[34] Gunnarsdottir, T., Njardvik, U., Olafsdottir, A.S., Craighead, L.W. and Bjarnason, R. (2011) Teasing and social rejection among obese children enrolling in family-based behavioural treatment: effects on psychological adjustment and academic competencies. *International Journal of Obesity 2012 36:1* **36**: 35–44. doi:10.1038/ijo.2011.181.

[35] Lydon, S., Healy, O., Roche, M., Henry, R., Mulhern, T. and Hughes, B.M. (2015) Salivary cortisol levels and challenging behavior in children with autism spectrum disorder. *Research in Autism Spectrum Disorders* **10**: 78–92. doi:10.1016/j.rasd.2014.10.020.

[36] Berrocal, A. and Wac, K. (2020) Peers know you: A feasibility study of the predictive value of peer's observations to estimate human states. *Procedia Computer Science* **175**: 205–213. doi:10.1016/j.procs.2020.07.031. The 17th International Conference on Mobile Systems and Pervasive Computing (MobiSPC),The 15th International Conference on Future Networks and Communications (FNC),The 10th International Conference on Sustainable Energy Information Technology.

[37] Berrocal, A., Concepcion, W., De Dominicis, S. and Wac, K. (2020) Complementing human behavior assessment by leveraging personal ubiquitous devices and social links: An evaluation of the peer-ceived momentary assessment method. *JMIR Mhealth Uhealth* **8**(8): e15947. doi:10.2196/15947.

[38] Schulz, P., Schlotz, W. and Becker, P. (2004) *TICS: Trierer Inventar zum chronischen Stress* (Hogrefe Verlag).

[39] Kovacs, M. (1992) Children's depression inventory manual. multi-health systems. *Inc. North Tonawanda, NY* .

[40] March, J.S., Parker, J.D., Sullivan, K., Stallings, P. and Conners, C.K. (1997) The multidimensional anxiety scale for children (masc): Factor structure, reliability, and validity. *Journal of the American Academy of Child and Adolescent Psychiatry* **36**: 554–565. doi:10.1097/00004583-199704000-00019.

[41] Goodman, R. (1997) The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry* **38**: 581–586. doi:10.1111/J.1469-7610.1997.TB01545.X.

[42] Groden, J., Diller, A., Bausman, M., Velicer, W., Norman, G. and Cautela, J. (2001) The development of a stress survey schedule for persons with autism and other developmental disabilities. *Journal of Autism and Developmental Disorders* **31**: 207–217. doi:10.1023/A:1010755300436/METRICS.

[43] Cohen, S., Kamarck, T. and Mermelstein, R. (1983) A global measure of perceived stress. *Journal of Health and Social Behavior* **24**(4): 385–396. URL http://www.jstor.org/stable/2136404.

[44] Crowne, D.P. and Marlowe, D. (1960) A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology* **24**: 349–354. doi:10.1037/H0047358.

[45] Holmes, T.H. and Rahe, R.H. (1967) The social readjustment rating scale. *Journal of Psychosomatic Research* **11**(2): 213–218. doi:10.1016/0022-3999(67)90010-4.

[46] Fanning, J. and McAuley, E. (2014) A comparison of tablet computer and paper-based questionnaires in healthy aging research. *JMIR Res Protoc* **3**(3): e38. doi:10.2196/resprot.3291.

[47] Mishra, P., Pandey, C., Singh, U., Keshri, A. and Sabaretnam, M. (2019) Selection of appropriate statistical methods for data analysis. *Annals of Cardiac Anaesthesia* **22**: 297–301. doi:10.4103/ACA.ACA_248_18.

[48] Xie, Y. (2011) Values and limitations of statistical models. *Research in Social Stratification and Mobility* **29**(3): 343–349. doi:10.1016/j.rssm.2011.04.001.

[49] Bzdok, D., Altman, N. and Krzywinski, M. (2018) Points of significance: Statistics versus machine learning. *Nature Methods* **15**: 233–234. doi:10.1038/NMETH.4642.